

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: EVANS, CHRISTOPHER J.
KEITH, DUANE E.

(ii) TITLE OF INVENTION: OPIOID RECEPTOR GENES

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: MORRISON & FOERSTER
(B) STREET: 2000 Pennsylvania Avenue, N.W. Suite 5500
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: USA
(F) ZIP: 20006-1888

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/403,260
(B) FILING DATE: 13-MAR-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: MURASHIGE, KATE H.
(B) REGISTRATION NUMBER: 29,959
(C) REFERENCE/DOCKET NUMBER: 22000-20526.21

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (202) 887-1500
(B) TELEFAX: (202) 887-0763

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Tyr Gly Gly Phe Met
1 5

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Tyr Gly Gly Phe Leu
1 5

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /note= "D form of amino acid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Tyr Ser Gly Phe Leu Thr
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: group(2, 5)
- (D) OTHER INFORMATION: /product= "OTHER"

/note= "D-penicillamine"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Tyr Xaa Gly Phe Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /note= "D form of amino acid"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /note= "D form of amino acid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Tyr Ala Gly Phe Leu
1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 2
- (D) OTHER INFORMATION: /note= "D form of amino acid"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 4
- (D) OTHER INFORMATION: /product= "MePhe"

/note= "N-Methylphenylalanine"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /product= "Gly-ol"

/note= "Carboxy end of glycine has been replaced with an alcohol substituent"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Tyr Ala Gly Xaa Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1829 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 29..1144

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCACGGTGGAA GACGGACACGG CGGGCGCC ATG GAG CTG GTG CCC TCT GCC CGT
Met Glu Leu Val Pro Ser Ala Arg
1 5

52

GCG GAG CTG CAG TCC TCG CCC CTC GTC AAC CTC TCG GAC GCC TTT CCC
Ala Glu Leu Gln Ser Ser Pro Leu Val Asn Leu Ser Asp Ala Phe Pro
10 15 20

100

AGC GCC TTC CCC AGC GCG GGC GCC AAT GCG TCG GGG TCG CCG GGA GCC
Ser Ala Phe Pro Ser Ala Gly Ala Asn Ala Ser Gly Ser Pro Gly Ala
25 30 35 40

148

CGT AGT GCC TCG TCC CTC GCC CTA GCC ATC GCC ATC ACC GCG CTC TAC Arg Ser Ala Ser Ser Leu Ala Leu Ala Ile Ala Ile Thr Ala Leu Tyr 45 50 55	196
TCG GCT GTG TGC GCA GTG GGG CTT CTG GGC AAC TGT CTC GTC ATG TTT Ser Ala Val Cys Ala Val Gly Leu Leu Gly Asn Cys Leu Val Met Phe 60 65 70	244
GGC ATC GTC CGG TAC ACC AAA TTG AAG ACC GCC ACC AAC ATC TAC ATC Gly Ile Val Arg Tyr Thr Lys Leu Lys Thr Ala Thr Asn Ile Tyr Ile 75 80 85	292
TTC AAT CTG GCT TTG GCT GAT GCG CTG GCC ACC AGC ACG CTG CCC TTC Phe Asn Leu Ala Leu Ala Asp Ala Leu Ala Thr Ser Thr Leu Pro Phe 90 95 100	340
CAG AGC GCC AAG TAC TTG ATG GAA ACG TGG CCG TTT GGC GAG CTG CTG Gln Ser Ala Lys Tyr Leu Met Glu Thr Trp Pro Phe Gly Glu Leu Leu 105 110 115 120	388
TGC AAG GCT GTG CTC TCC ATT GAC TAC TAC AAC ATG TTC ACT AGC ATC Cys Lys Ala Val Leu Ser Ile Asp Tyr Tyr Asn Met Phe Thr Ser Ile 125 130 135	436
TTC ACC CTC ACC ATG ATG AGC GTG GAC CGC TAC ATT GCT GTC TGC CAT Phe Thr Leu Thr Met Met Ser Val Asp Arg Tyr Ile Ala Val Cys His 140 145 150	484
CCT GTC AAA GCC CTG GAC TTC CGG ACA CCA GCC AAG GCC AAG CTG ATC Pro Val Lys Ala Leu Asp Phe Arg Thr Pro Ala Lys Ala Lys Leu Ile 155 160 165	532
AAT ATA TGC ATC TGG GTC TTG GCT TCA GGT GTC GGG GTC CCC ATC ATG Asn Ile Cys Ile Trp Val Leu Ala Ser Gly Val Gly Val Pro Ile Met 170 175 180	580
GTC ATG GCA GTG ACC CAA CCC CGG GAT GGT GCA GTG GTA TGC ATG CTC Val Met Ala Val Thr Gln Pro Arg Asp Gly Ala Val Val Cys Met Leu 185 190 195 200	628
CAG TTC CCC AGT CCC AGC TGG TAC TGG GAC ACT GTG ACC AAG ATC TGC Gln Phe Pro Ser Pro Ser Trp Tyr Trp Asp Thr Val Thr Lys Ile Cys 205 210 215	676
GTG TTC CTC TTT GCC TTC GTG GTG CCG ATC CTC ATC ATC ACG GTG TGC Val Phe Leu Phe Ala Phe Val Val Pro Ile Leu Ile Ile Thr Val Cys 220 225 230	724
TAT GGC CTC ATG CTA CTG CGC CTG CGC AGC GTG CGT CTG CTG TCC GGT Tyr Gly Leu Met Leu Leu Arg Leu Arg Ser Val Arg Leu Leu Ser Gly 235 240 245	772
TCC AAG GAG AAG GAC CGC AGC CTG CGG CGC ATC ACG CGC ATG GTG CTG Ser Lys Glu Lys Asp Arg Ser Leu Arg Arg Ile Thr Arg Met Val Leu 250 255 260	820
GTG GTG GTG GGC GCC TTC GTG GTG TGC TGG GCG CCC ATC CAC ATC TTC Val Val Val Gly Ala Phe Val Val Cys Trp Ala Pro Ile His Ile Phe 265 270 275 280	868
GTC ATC GTC TGG ACG CTG GTG GAC ATC AAT CGG CGC GAC CCA CTT GTG Val Ile Val Trp Thr Leu Val Asp Ile Asn Arg Arg Asp Pro Leu Val 285 290 295	916

GTG GCC GCA CTG CAC CTG TGC ATT GCG CTG GGC TAC GCC AAC AGC AGC Val Ala Ala Leu His Leu Cys Ile Ala Leu Gly Tyr Ala Asn Ser Ser 300 305 310	964
CTC AAC CCG GTT CTC TAC GCC TTC CTG GAC GAG AAC TTC AAG CGC TGC Leu Asn Pro Val Leu Tyr Ala Phe Leu Asp Glu Asn Phe Lys Arg Cys 315 320 325	1012
TTC CGC CAG CTC TGT CGC ACG CCC TGC GGC CGC CAA GAA CCC GGC AGT Phe Arg Gln Leu Cys Arg Thr Pro Cys Gly Arg Gln Glu Pro Gly Ser 330 335 340	1060
CTC CGT CGT CCC CGC CAG GCC ACC ACG CGT GAG CGT GTC ACT GCC TGC Leu Arg Arg Pro Arg Gln Ala Thr Thr Arg Glu Arg Val Thr Ala Cys 345 350 355 360	1108
ACC CCC TCC GAC GGC CCG GGC GGT GGC GCT GCC GCC TGACCTACCC Thr Pro Ser Asp Gly Pro Gly Gly Ala Ala Ala 365 370	1154
GACCTTCCCC TAAACGCC CTCCCAAGTG AAGTGATCAG AGGCCACACC GAGCTCCCTG GGAGGCTGTG GCCACCACCA GGACAGCTAG AATTGGGCCT GCACAGAGGG GAGGCCTCCT GTGGGACGG GCCTGAGGGA TCAAAGGCTC CAGGTTGGAA CGGTGGGGGT GAGGAAGCAG AGCTGGTGAT TCCTAAACTG TATCCATTAG TAAGGCCTCT CAATGGGACA GAGCCTCCGC CTTGAGATAA CATCGGGTTC TGGCCTTTT GAACACCCAG CTCCAGTCCA AGACCCAAGG ATTCCAGCTC CAGAACCAAGG AGGGGCAGTG ATGGGGTCGA TGATTTGGTT TGGCTGAGAG TCCCAGCATT TGTGTTATGG GGAGGATCTC TCATCTTAGA GAAGAAAGGG GACAGGGCAT TCAGGCAAGG CAGCTTGGGG TTTGGTCAGG AGATAAGCGC CCCCCTTCCC TTGGGGGGAG GATAAGTGGG GGATGGTCAC GTGGAGAAG AGTCAAAGTT CTCACCACCT TTCTAACTAC TCAGCTAAC TCGTTGAGGC TAGGGCAAC GTGACTTCTC TGTAGAGAGG TACAAGCCGG GCCTGATGGG GCAGGCCTGT GTAATCCCAG TCATAGTGGA GGCTGAGGCT GGAAAATTAA GGACCAACAG CCCGG	1214 1274 1334 1394 1454 1514 1574 1634 1694 1754 1814 1829

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Glu Leu Val Pro Ser Ala Arg Ala Glu Leu Gln Ser Ser Pro Leu 1 5 10 15
Val Asn Leu Ser Asp Ala Phe Pro Ser Ala Phe Pro Ser Ala Gly Ala 20 25 30
Asn Ala Ser Gly Ser Pro Gly Ala Arg Ser Ala Ser Ser Leu Ala Leu 35 40 45

Ala Ile Ala Ile Thr Ala Leu Tyr Ser Ala Val Cys Ala Val Gly Leu
50 55 60

Leu Gly Asn Cys Leu Val Met Phe Gly Ile Val Arg Tyr Thr Lys Leu
65 70 75 80

Lys Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu Ala Leu Ala Asp Ala
85 90 95

Leu Ala Thr Ser Thr Leu Pro Phe Gln Ser Ala Lys Tyr Leu Met Glu
100 105 110

Thr Trp Pro Phe Gly Glu Leu Leu Cys Lys Ala Val Leu Ser Ile Asp
115 120 125

Tyr Tyr Asn Met Phe Thr Ser Ile Phe Thr Leu Thr Met Met Ser Val
130 135 140

Asp Arg Tyr Ile Ala Val Cys His Pro Val Lys Ala Leu Asp Phe Arg
145 150 155 160

Thr Pro Ala Lys Ala Lys Leu Ile Asn Ile Cys Ile Trp Val Leu Ala
165 170 175

Ser Gly Val Gly Val Pro Ile Met Val Met Ala Val Thr Gln Pro Arg
180 185 190

Asp Gly Ala Val Val Cys Met Leu Gln Phe Pro Ser Pro Ser Trp Tyr
195 200 205

Trp Asp Thr Val Thr Lys Ile Cys Val Phe Leu Phe Ala Phe Val Val
210 215 220

Pro Ile Leu Ile Ile Thr Val Cys Tyr Gly Leu Met Leu Leu Arg Leu
225 230 235 240

Arg Ser Val Arg Leu Leu Ser Gly Ser Lys Glu Lys Asp Arg Ser Leu
245 250 255

Arg Arg Ile Thr Arg Met Val Leu Val Val Val Gly Ala Phe Val Val
260 265 270

Cys Trp Ala Pro Ile His Ile Phe Val Ile Val Trp Thr Leu Val Asp
275 280 285

Ile Asn Arg Arg Asp Pro Leu Val Val Ala Ala Leu His Leu Cys Ile
290 295 300

Ala Leu Gly Tyr Ala Asn Ser Ser Leu Asn Pro Val Leu Tyr Ala Phe
305 310 315 320

Leu Asp Glu Asn Phe Lys Arg Cys Phe Arg Gln Leu Cys Arg Thr Pro
325 330 335

Cys Gly Arg Gln Glu Pro Gly Ser Leu Arg Arg Pro Arg Gln Ala Thr
340 345 350

Thr Arg Glu Arg Val Thr Ala Cys Thr Pro Ser Asp Gly Pro Gly Gly
355 360 365

Gly Ala Ala Ala
370

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 369 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met	Glu	Leu	Thr	Ser	Glu	Gln	Phe	Asn	Gly	Ser	Gln	Val	Trp	Ile	Pro
1					5					10				15	
Ser	Pro	Phe	Asp	Leu	Asn	Gly	Ser	Leu	Gly	Pro	Ser	Asn	Gly	Ser	Asn
				20				25					30		
Gln	Thr	Glu	Pro	Tyr	Tyr	Asp	Met	Thr	Ser	Asn	Ala	Val	Leu	Thr	Phe
				35				40					45		
Ile	Tyr	Phe	Val	Val	Cys	Val	Val	Gly	Leu	Cys	Gly	Asn	Thr	Leu	Val
				50				55			60				
Ile	Tyr	Val	Ile	Leu	Arg	Tyr	Ala	Lys	Met	Lys	Thr	Ile	Thr	Asn	Ile
				65				70		75			80		
Tyr	Ile	Leu	Asn	Leu	Ala	Ile	Ala	Asp	Glu	Leu	Phe	Met	Leu	Gly	Leu
				85				90					95		
Pro	Phe	Leu	Ala	Met	Gln	Val	Ala	Leu	Val	His	Trp	Pro	Phe	Gly	Lys
				100				105					110		
Ala	Ile	Cys	Arg	Val	Val	Met	Thr	Val	Asp	Gly	Ile	Asn	Gln	Phe	Thr
				115				120				125			
Ser	Ile	Phe	Cys	Leu	Thr	Val	Met	Ser	Ile	Asp	Arg	Tyr	Leu	Ala	Val
				130				135			140				
Val	His	Pro	Ile	Lys	Ser	Ala	Lys	Trp	Arg	Arg	Pro	Arg	Thr	Ala	Lys
				145				150		155			160		
Met	Ile	Asn	Val	Ala	Val	Trp	Gly	Val	Ser	Leu	Leu	Val	Ile	Leu	Pro
				165				170					175		
Ile	Met	Ile	Tyr	Ala	Gly	Leu	Arg	Ser	Asn	Gln	Trp	Gly	Arg	Ser	Ser
				180				185				190			
Cys	Thr	Ile	Asn	Trp	Pro	Gly	Glu	Ser	Gly	Ala	Trp	Tyr	Thr	Gly	Phe
				195				200				205			
Ile	Ile	Tyr	Ala	Phe	Ile	Leu	Gly	Phe	Leu	Val	Pro	Leu	Thr	Ile	Ile
				210				215				220			
Cys	Leu	Cys	Tyr	Leu	Phe	Ile	Ile	Ile	Lys	Val	Lys	Ser	Ser	Gly	Ile
				225				230			235			240	
Arg	Val	Gly	Ser	Ser	Lys	Arg	Lys	Lys	Ser	Glu	Lys	Lys	Val	Thr	Arg
				245				250					255		
Met	Val	Ser	Ile	Val	Val	Ala	Val	Phe	Ile	Phe	Cys	Trp	Leu	Pro	Phe
				260				265					270		
Tyr	Ile	Phe	Asn	Val	Ser	Ser	Val	Ser	Val	Ala	Ile	Ser	Pro	Thr	Pro
				275				280				285			
Ala	Leu	Lys	Gly	Met	Phe	Asp	Phe	Val	Val	Ile	Leu	Thr	Tyr	Ala	Asn
				290				295			300				

Ser Cys Ala Asn Pro Ile Leu Tyr Ala Phe Leu Ser Asp Asn Phe Lys
305 310 315 320

Lys Ser Phe Gln Asn Val Leu Cys Leu Val Lys Val Ser Gly Ala Glu
325 330 335

Asp Gly Glu Arg Ser Asp Ser Lys Gln Asp Lys Ser Arg Leu Asn Glu
340 345 350

Thr Thr Glu Thr Gln Arg Thr Leu Leu Asn Gly Asp Leu Gln Thr Ser
355 360 365

Ile

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGCAGTGGT GTGCATGCTC CAGTTCCCCA GCCCCAGCTG GTACTGGGAC ACGGTGACCA 60
AGATCTGCGT GTTCCTCTTC GCCTTCGTGG TGCCCATCCT CATCATCACCC GTGTGCTATG 120
GCCTCATGCT 130

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGTGCAGTGG TATGCATGCT CCAGTTCCCC AGTCCCAGCT GGTACTGGGA CACTGTGACC 60
AAGATCTGCG TGTTCCCTCTT TGCCTTCGTG GTGCCGATCC TCATCATCAC GGTGTGCTAT 120
GGCCTCATGC 130

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2447 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCTGGCCTTT TGGGGATGTG CTGTGCAAGA TAGTAATTTC CATTGATTAC TACAACATGT 60
TCACCAAGCAT CTTCACCTTG ACCATGATGA GCGTGGACCG CTACATTGCC GTGTGCCACC 120
CCGTGAAGGC TTTGGACTTC CGCACACCCT TGAAGGCAGA GATCATCAAT ATCTGCATCT 180

GGCTGCTGTC	GTCATCTGTT	GGCATCTCTG	CAATAGTCCT	TGGAGGCACC	AAAGTCAGGG	240
AAGGTAAGAG	CAGTCATTTC	ATTCTGTTCA	TAAAAATGTA	GCTTCAAATT	ACATAGACTT	300
TTAATTGAG	CGTGAGTAGG	CCACATATTT	GTGGAAATCG	ATGCCAAAAG	ACGACGGAAA	360
TGTAGTGCCT	AAATCCATGG	AAGATGAGAA	GTAGAACAAAT	TTTTGTCCC	TTTCCACCTC	420
TAAACACAGA	ATGCAATAAT	GACATTGCCA	GAAGAGAGAT	GCCCGACCTG	TCTCCCATT	480
TGGCAATGTT	TAGTAGAAAG	TGGAGGGGTG	AGGATGAGGT	AAGAACACAA	GGCATGTAGA	540
TTTTAAAGTA	CAACCTGGCA	AGTCCAGACA	CACCTTCTCA	CTCCTTTTT	TCTCTTTAAC	600
AAGGGATATA	AATTATTGGT	GACATATGCT	GGTTGTTCC	TCTTTTATT	CTAAAGGATA	660
ACCTCCAAAT	CACTATTTA	ACAGCTTGG	CGTAGGATCT	CAAATCAAG	TTAACGGATG	720
GTAGTTACAG	ATGAGTCAGA	ACCACTTGAT	TTGGACATAT	CAGGTTTCC	CTTGCAAACC	780
AGCCAACGTGA	TTTTTTTTT	TTTTTTTTT	GAGAGAGAGT	CTTGCTCTGT	TGCCAGGCTA	840
GAGTGCAGTG	GCGCGATATC	GGCTCACTGC	AACCTCTGCC	TCCCGGGTTC	AACCTCAGCC	900
TCTCGAGTAG	CTGGGACTAC	TGGCACACAC	CACCATGCC	AGCTAATT	TGTATT	960
GTAGAGACAG	GGTTTCACCG	TGTTGCCAG	GGTGGTCTCA	ATCTCTTGAC	CTCGTGATCT	1020
GCCCGCCTCG	NCTCCCCAAA	GTGCTGGAT	TACAGGCGTG	CNCTGCNCCC	GNCCCCTGTT	1080
GATGTTTTC	CTGTATTCT	AGGACAGTAG	TTCTCACTCT	GGGCTGCACA	TTGGAATCAC	1140
CTGGGTACTT	TAGAAAACAC	TGCTGCCTGC	ATCCCACCCC	TTAAGGGTCT	GGTGTAAATTG	1200
ACCTGGGGTA	CAGCCTGGGT	GTCAAGATT	TTGAGCTCTC	TCCAGGTGAC	TCTGACCTGC	1260
AGCCAAGGTG	AGAGGTACTG	TTCTAGGAGT	TTGCTTTAC	TAGCAAAATA	TAAAGCTATA	1320
GAAAGCATCT	TTTGTTCCTC	ATAGAAATT	ATGATGGGA	GGTGAGCAGA	ATAGTCACTC	1380
TGGGCCTACT	CATGCTGTT	AATGCTCCAG	CAGGTATATA	GGTCTCCAG	TTACTAGGG	1440
GTTCATAATA	CCTGTGAGAG	CAGATAACTG	AGTGTATATA	GTGAGGATT	CCAGGTCTATA	1500
GTGAAAGGGC	AAGGCACTAA	AATCATAGCT	TGTCTTGAT	ATACTGTTG	TTGTTTTA	1560
GACTTACATG	TTAGGTTCA	GTTTACGTT	TAGGTTACA	GCAAAACTGA	CCAGAAAGCA	1620
CAGAGAGGCA	CTTCNATT	CCTCCATT	CCCCACACAG	GCACATCCTC	CCCTACAGAG	1680
TGGTCCATT	ATTACAGCTG	CTGAACCCAC	ACTGACACGC	TGTTATCACT	CAGAGCCTGG	1740
CAGTTACAG	AGGCTCACTC	TCCGNTATGT	GTCCTGTGNT	TTGAACAAAT	GTATAATGAC	1800
TTTATTCTATT	GTGTTTTAAT	GAAGCTGATC	TTTCCCTCT	GAAACTACAA	AATGAATT	1860
TAGCATAGCC	ATAGCAGGTG	TCAAGCTATA	CTACTAGGTA	AATTAAAGA	AATGCCAAC	1920
TTTATCATAT	TTGCATTCA	AAATATGATT	AATCACACAT	AGGATTTGT	TTCTTCATGC	1980
CTACAGCAAA	TAGAAATAAA	GTGCAAGAAA	CTTTCTGAG	GCAAGCTTT	CACTTTGTGA	2040
ACGTAAAATG	TTGACTCTAA	TATTTCCAT	ACTGTAGTAT	ATGTGTGTGT	ATTATGTGAG	2100
GATTCAAGT	CTGCTCTTAC	TTTTTATAG	TAGCTAAGAA	TTATTATAAT	CGCTATAAGC	2160

AGAAACAATT ATTCTTAACA AAATGAATAAC ACACAAGAAA AGCTTAGTT TAGCTATTAG	2220
AACTAACTCT ATAATTATGA TAACCAGAG ATGCTGGAAC AGGAGCCAGC AGAAGCCACA	2280
GCCCTCTGAT ATTAATATAT AAAGAAACCA AAATCTGCTT GTTAAACTGA GGCAGTTGTA	2340
TGGATACTTC AACCTGAAAA TGCCCCCTTC TTCCTGAAAC AGAACATTAA ATAAAAATGG	2400
CATGCTTGGGA CAGGAATTTC TTTTTAAAAA AATGCTTAGT TTTTATG	2447

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 830 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTCCCTTATC TCCTAGATAC ACCAAGATGA AGACTGCCAC CAACATCTAC ATTTCAACC	60
TTGCTCTGCA GATGCCTTAG CCACCAAGTAC CCTGCCCTTC CAGAGTGTGA ATTACCTAAT	120
GGGAACATGG CCATTTGGAA CCATCCTTG CAAGATAGTG ATCTCCATAG ATTACTATAA	180
CATGTTCACCC AGCATATTCA CCCTCTGCAC CATGAGTGTGTT GATCGATACA TTGCAGTCTG	240
CCACCCCTGTC AAGGCCTTAG ATTTCCGTAC TCCCCNNNN NNNNNNNNNNN NNNNNNNNNNN	300
NNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN	360
NNNNNNNNNGT TCCATAGATT GTACACTAAC ATTCTCTCAT CCAACCTGGT ACTGGGAAAA	420
CCTGCTGAAG ATCTGTGTTT TCATCTTCGC CTTCAATTATG CCAGTGCTCA TCATTACCGT	480
GTGCTATGGA CTGATGATCT TGCGCCTCAA GAGTGTCCGC ATGCTCTCTG GCTCCAAAGA	540
AAAGGACAGG AATCTTCGAA GGATCACCAAG GATGGTGCTG GTGGTGGTGG CTGTGTTCAT	600
CGTCTGCTGG ACTCCCATTAC ACATTACGT CATCATTAAA GCCTTGGTTA CAATCCCAGA	660
AACTACGTTTC CAGACTGTTT CTTGGCACTT CTGCATTGCT CTAGGTTACA CAAACAGCTG	720
CCTCAACCCA GTCCCTTATG CATTCTGGA TGAAAACCTTC CACGATGCTT CAGAGAGTTC	780
TGTATCCCAA CCTCTTCCAA CATTGAGCAA CAAAACCTCCA CTCGAATTCC	830

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 332 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGGTACCGGG CCCCCCCCTCG AGGTCGACGG TATCGATAAG CTTGATATCG AATTCTTACT	60
GAATTAGGTA TCTTCTTCA CACTACTTGG TAAAAAAAT GAAAAGGCAG AAAAATTAGC	120
CCCAAAAGAG ATGAAAACCTCT TCCGTCCATC ACCATTGACT CTATTGTGAA CTTATGAAAA	180

AGGTAGTTGA GCAATATGAA GGCCATGATG TGGAATTAAA CACACACACA CACACACACA	240
CACACACACA CACATGCTGG ATTCTAAATG TGTCCCTCCT CCTCTCACTC TCTTGATTCA	300
AGTTTATTTC TGAAGTGAGA CACGATCACC AC	332

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1981 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGGATCCTTA GCATCCCCAA AGCGCCTCCG TGTACTTCTA AGGTGGGAGG GGGATACAAG	60
CAGAGGAGAA TATCGGACGC TCAGACGTT CATTCTGCCT GCCGCTCTTC TCTGGTTCCA	120
CTAGGGCTTG TCCTTGTAAG AAACTGACGG AGCCTAGGGC AGCTGTGAGA GGAAGAGGCT	180
GGGGCGCCTG GAACCCGAAC ACTCTTGAGT GCTCTCAGTT ACAGNCTACC GAGTCCGCAG	240
GAAGCATTCA GAACCATGGA CAGCAGCGCC GGCCCAGGGA ACATCAGCGA CTGCTCTGAC	300
CCCTTAGCTC CTGCAAGTTG CTCCCCAGCA CCTGGCTCCT GGCTCAACTT GTCCCACGTT	360
GATGGAAACC AGTCCGACCC ATGCGGTCC AACCCGACGG GCCTTGGCGG GAACGACAGC	420
CTGTGCCCTC AGACCGGCAG CCCTTCCATG GTCACAGCCA TCACCATCAT GGCCCTCTAT	480
TCTATCGTGT GTGTAGTGGG CCTCTTGGA AACTTCCTGG TCATGTATGT GATTGTAAGA	540
TATACCAAAA TGAAGACTGC CACCAACATC TACATTTCA ACCTTGCTCT GGCAGATGCC	600
TTAGCCACTA GCACGCTGCC CTTTCAGAGT GTTAACTACC TGATGGGAAC GTGGCCCTTT	660
GGAAACATCC TCTGCAAGAT CGTGATCTCA ATAGACTACT ACAACATGTT CACCAGTATC	720
TTCACCCCTCT GCACCATGAG TGTAGACCGC TACATTGCCG TCTGCCACCC GGTCAAGGCC	780
CTGGATTTCG GTACCCCCCG AAATGCCAAA ATTGTCAATG TCTGCAACTG GATCCTCTCT	840
TCTGCCATTG GTCTGCCCGT AATGTTCATG GCAACCACAA AATACAGGCA GGGGTCCATA	900
GATTGCACCC TCACGTTCTC TCATCCCACA TGGTACTGGG AGAACCTGCT CAAAATCTGT	960
GTCTTCATCT TCGCCTTCAT CATGCCGGC CTCATCATCA CTGTGTGTTA TGGACTGATG	1020
ATCTTACAGC TCAAGAGTGT CCGCATGCTG TCGGGCTCCA AAGAAAAGGA CAGGAACCTG	1080
CGCAGGATCA CCCGGATGGT GCTGGTGGTC GTGGCTGTAT TTATTGTCTG CTGGACCCCC	1140
ATCCACATCT ATGTCATCAT CAAAGCACTG ATCACGATTG CAGAAACCAC TTTCCAGACT	1200
GTTTCTGGC ACTTCTGCAT TGCCTTGGGT TACACAAACA GCTGCCGTGAA CCCAGTTCTT	1260
TATGCGTTCC TGGATGAAAA CTTCAAACGA TGTTTAGAG AGTTCTGCAT CCCAACTTCC	1320
TCCACAATCG AACAGCAAAA CTCTGCTCGA ATCCGTCAAACACACTAGGGAA ACACCCCTCC	1380
ACGGCTAATA CAGTGGATCG AACTAACCAC CAGCTAGAAA ATCTGGAAGC AGAAACTGCT	1440
CCATTGCCCT AACTGGGTCC CACGCCATCC AGACCCTCGC TAAACTTAGA GGCTGCCATC	1500

TACCTTCCAAT CAGGTTGCTG TCAGGGTTTG TGGGAGGCTC TGGTTTCCTG GAAAAGCATC	1560
TGATCCTGCA TCATTCAAAG TCATTCTCT CTGGCTATTG ACGCTACACG TCAGAGACAC	1620
TCAGACTGTG TCAAGCACTC AGAAGGAAGA GACTGCAGGC CACTACTGAA TCCAGCTCAT	1680
GTACAGAAC ATCCAATGGA CCACAATACT CTGTGGTATG TGATTTGTGA TCAACATAGA	1740
AGGTGACCCCT TCCCTATGTG GAATTAA TTTCAAGGAA ATACTTATGA TCTCATCAAG	1800
GGAAAAATAG ATGTCACTTG TTAAATTAC TGTAGTGATG CATAAAGGAA AAGCTACCTC	1860
TGACCTCTAG CCCAGTCACC CTCTATGGAA AGTCCATAG GGAATATGTG AGGGAAAATG	1920
TTGCTTCCAA ATTAAATTAC CACCTTATG TTATAGTCTA GTTAAGACAT CAGGGGCATC	1980
T	1981

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 398 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Asp Ser Ser Thr Gly Pro Gly Asn Thr Ser Asp Cys Ser Asp Pro			
1	5	10	15
Leu Ala Gln Ala Ser Cys Ser Pro Ala Pro Gly Ser Trp Leu Asn Leu			
20	25	30	
Ser His Val Asp Gly Asn Gln Ser Asp Pro Cys Gly Leu Asn Arg Thr			
35	40	45	
Gly Leu Gly Gly Asn Asp Ser Leu Cys Pro Gln Thr Gly Ser Pro Ser			
50	55	60	
Met Val Thr Ala Ile Thr Ile Met Ala Leu Tyr Ser Ile Val Cys Val			
65	70	75	80
Val Gly Leu Phe Gly Asn Phe Leu Val Met Tyr Val Ile Val Arg Tyr			
85	90	95	
Thr Lys Met Lys Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu Ala Leu			
100	105	110	
Ala Asp Ala Leu Ala Thr Ser Thr Leu Pro Phe Gln Ser Val Asn Tyr			
115	120	125	
Leu Met Gly Thr Trp Pro Phe Gly Thr Ile Leu Cys Lys Ile Val Ile			
130	135	140	
Ser Ile Asp Tyr Tyr Asn Met Phe Thr Ser Ile Phe Thr Leu Cys Thr			
145	150	155	160
Met Ser Val Asp Arg Tyr Ile Ala Val Cys His Pro Val Lys Ala Leu			
165	170	175	
Asp Phe Arg Thr Pro Arg Asn Ala Lys Ile Val Asn Val Cys Asn Trp			
180	185	190	

Ile Leu Ser Ser Ala Ile Gly Leu Pro Val Met Phe Met Ala Thr Thr
195 200 205
Lys Tyr Arg Gln Gly Ser Ile Asp Cys Thr Leu Thr Phe Ser His Pro
210 215 220
Thr Trp Tyr Trp Glu Asn Leu Leu Lys Ile Cys Val Phe Ile Phe Ala
225 230 235 240
Phe Ile Met Pro Ile Leu Ile Thr Val Cys Tyr Gly Leu Met Ile
245 250 255
Leu Arg Leu Lys Ser Val Arg Met Leu Ser Gly Ser Lys Glu Lys Asp
260 265 270
Arg Asn Leu Arg Arg Ile Thr Arg Met Val Leu Val Val Val Ala Val
275 280 285
Phe Ile Val Cys Trp Thr Pro Ile His Ile Tyr Val Ile Ile Lys Ala
290 295 300
Leu Ile Thr Ile Pro Glu Thr Thr Phe Gln Thr Val Ser Trp His Phe
305 310 315 320
Cys Ile Ala Leu Gly Tyr Thr Asn Ser Cys Leu Asn Pro Val Leu Tyr
325 330 335
Ala Phe Leu Asp Glu Asn Phe Lys Arg Cys Phe Arg Glu Phe Cys Ile
340 345 350
Pro Thr Ser Ser Thr Ile Glu Gln Gln Asn Ser Thr Arg Val Arg Gln
355 360 365
Asn Thr Arg Glu His Pro Ser Thr Ala Asn Thr Val Asp Arg Thr Asn
370 375 380
His Gln Leu Glu Asn Leu Glu Ala Glu Thr Ala Pro Leu Pro
385 390 395

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Glu Ser Pro Ile Gln Ile Phe Arg Gly Asp Pro Gly Pro Thr Cys
1 5 10 15
Ser Pro Ser Ala Cys Leu Leu Pro Asn Ser Ser Ser Trp Phe Pro Asn
20 25 30
Trp Ala Glu Ser Asp Ser Asn Gly Ser Val Gly Ser Glu Asp Gln Gln
35 40 45
Leu Glu Ser Ala His Ile Ser Pro Ala Ile Pro Val Ile Ile Thr Ala
50 55 60
Val Tyr Ser Val Val Phe Val Val Gly Leu Val Gly Asn Ser Leu Val
65 70 75 80

Met Phe Val Ile Ile Arg Tyr Thr Lys Met Lys Thr Ala Thr Asn Ile
85 90 95

Tyr Ile Phe Asn Leu Ala Leu Ala Asp Ala Leu Val Thr Thr Thr Met
100 105 110

Pro Phe Gln Ser Ala Val Tyr Leu Met Asn Ser Trp Pro Phe Gly Asp
115 120 125

Val Leu Cys Lys Ile Val Ile Ser Ile Asp Tyr Tyr Asn Met Phe Thr
130 135 140

Ser Ile Phe Thr Leu Thr Met Met Ser Val Asp Arg Tyr Ile Ala Val
145 150 155 160

Cys His Pro Val Lys Ala Leu Asp Phe Arg Thr Pro Leu Lys Ala Lys
165 170 175

Ile Ile Asn Ile Cys Ile Trp Leu Leu Ala Ser Ser Val Gly Ile Ser
180 185 190

Ala Ile Val Leu Gly Gly Thr Lys Val Arg Glu Asp Val Ile Glu Cys
195 200 205

Ser Leu Gln Phe Pro Asp Asp Glu Trp Trp Asp Leu Phe Met Lys Ile
210 215 220

Cys Val Phe Val Phe Ala Phe Val Ile Pro Val Leu Ile Ile Ile Val
225 230 235 240

Cys Tyr Thr Leu Met Ile Leu Arg Leu Lys Ser Val Arg Leu Leu Ser
245 250 255

Gly Ser Arg Glu Lys Asp Arg Asn Leu Arg Arg Ile Thr Lys Leu Val
260 265 270

Leu Val Val Val Ala Val Phe Ile Ile Cys Trp Thr Pro Ile His Ile
275 280 285

Phe Ile Leu Val Glu Ala Leu Gly Ser Thr Ser His Ser Thr Ala Ala
290 295 300

Leu Ser Ser Tyr Tyr Phe Cys Ile Ala Leu Gly Tyr Thr Asn Ser Ser
305 310 315 320

Leu Asn Pro Val Leu Tyr Ala Phe Leu Asp Glu Asn Phe Lys Arg Cys
325 330 335

Phe Arg Asp Phe Cys Phe Pro Ile Lys Met Arg Met Glu Arg Gln Ser
340 345 350

Thr Asn Arg Val Arg Asn Thr Val Gln Asp Pro Ala Ser Met Arg Asp
355 360 365

Val Gly Gly Met Asn Lys Pro Val
370 375

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1805 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 10..1119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGCAGTGGC ATG GAG CCC CTC TTC CCC GCG CCG TTC TGG GAG GTT ATC	48
Met Glu Pro Leu Phe Pro Ala Pro Phe Trp Glu Val Ile	
375 380 385	
TAC GGC AGC CAC CTT CAG GGC AAC CTG TCC CTC CTG AGC CCC AAC CAC	96
Tyr Gly Ser His Leu Gln Gly Asn Leu Ser Leu Leu Ser Pro Asn His	
390 395 400	
AGT CTG CTG CCC CCG CAT CTG CTG CTC AAT GCC AGC CAC GGC GCC TTC	144
Ser Leu Leu Pro Pro His Leu Leu Leu Asn Ala Ser His Gly Ala Phe	
405 410 415	
CTG CCC CTC GGG CTC AAG GTC ACC ATC GTG GGG CTC TAC CTG GCC GTG	192
Leu Pro Leu Gly Leu Lys Val Thr Ile Val Gly Leu Tyr Leu Ala Val	
420 425 430	
TGT GTC GGA GGG CTC CTG GGG AAC TGC CTT GTC ATG TAC GTC ATC CTC	240
Cys Val Gly Leu Leu Gly Asn Cys Leu Val Met Tyr Val Ile Leu	
435 440 445	
AGG CAC ACC AAA ATG AAG ACA GCC ACC AAT ATT TAC ATC TTT AAC CTG	288
Arg His Thr Lys Met Lys Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu	
450 455 460 465	
GCC CTG GCC GAC ACT CTG GTC CTG CTG ACG CTG CCC TTC CAG GGC ACG	336
Ala Leu Ala Asp Thr Leu Val Leu Leu Thr Leu Pro Phe Gln Gly Thr	
470 475 480	
GAC ATC CTC CTG GGC TTC TGG CCG TTT GGG AAT GCG CTG TGC AAG ACA	384
Asp Ile Leu Leu Gly Phe Trp Pro Phe Gly Asn Ala Leu Cys Lys Thr	
485 490 495	
GTC ATT GCC ATT GAC TAC AAC ATG TTC ACC AGC ACC TTC ACC CTA	432
Val Ile Ala Ile Asp Tyr Tyr Asn Met Phe Thr Ser Thr Phe Thr Leu	
500 505 510	
ACT GCC ATG AGT GTG GAT CGC TAT GTA GCC ATC TGC CAC CCC ATC CGT	480
Thr Ala Met Ser Val Asp Arg Tyr Val Ala Ile Cys His Pro Ile Arg	
515 520 525	
GCC CTC GAC GTC CGC ACG TCC AGC AAA GCC CAG GCT GTC AAT GTG GCC	528
Ala Leu Asp Val Arg Thr Ser Ser Lys Ala Gln Ala Val Asn Val Ala	
530 535 540 545	
ATC TGG GCC CTG GCC TCT GTT GTC GGT GTT CCC GTT GCC ATC ATG GGC	576
Ile Trp Ala Leu Ala Ser Val Val Gly Val Pro Val Ala Ile Met Gly	
550 555 560	
TCG GCA CAG GTC GAG GAT GAA GAG ATC GAG TGC CTG GTG GAG ATC CCT	624
Ser Ala Gln Val Glu Asp Glu Glu Ile Glu Cys Leu Val Glu Ile Pro	
565 570 575	
ACC CCT CAG GAT TAC TGG GGC CCG GTG TTT GCC ATC TGC ATC TTC CTC	672
Thr Pro Gln Asp Tyr Trp Gly Pro Val Phe Ala Ile Cys Ile Phe Leu	
580 585 590	
TTC TCC TTC ATC GTC CCC GTG CTC GTC ATC TCT GTC TGC TAC AGC CTC	720
Phe Ser Phe Ile Val Pro Val Leu Val Ile Ser Val Cys Tyr Ser Leu	
595 600 605	

ATG ATC CGG CGG CTC CGT GGA GTC CGC CTG CTC TCG GGC TCC CGA GAG Met Ile Arg Arg Leu Arg Gly Val Arg Leu Leu Ser Gly Ser Arg Glu 610 615 620 625	768
AAG GAC CGG AAC CTG CGG CGC ATC ACT CGG CTG GTG CTG GTG GTA GTG Lys Asp Arg Asn Leu Arg Arg Ile Thr Arg Leu Val Leu Val Val Val 630 635 640	816
GCT GTG TTC GTG GGC TGC TGG ACG CCT GTC CAG GTC TTC GTG CTG GCC Ala Val Phe Val Gly Cys Trp Thr Pro Val Gln Val Phe Val Leu Ala 645 650 655	864
CAA GGG CTG GGG GTT CAG CCG AGC AGC GAG ACT GCC GTG GCC ATT CTG Gln Gly Leu Gly Val Gln Pro Ser Ser Glu Thr Ala Val Ala Ile Leu 660 665 670	912
CGC TTC TGC ACG GCC CTG GGC TAC GTC AAC AGC TGC CTC AAC CCC ATC Arg Phe Cys Thr Ala Leu Gly Tyr Val Asn Ser Cys Leu Asn Pro Ile 675 680 685	960
CTC TAC GCC TTC CTG GAT GAG AAC TTC AAG GCC TGC TTC CGC AAG TTC Leu Tyr Ala Phe Leu Asp Glu Asn Phe Lys Ala Cys Phe Arg Lys Phe 690 695 700 705	1008
TGC TGT GCA TCT GCC CTG CGC CGG GAC GTG CAG GTG TCT GAC CGC GTG Cys Cys Ala Ser Ala Leu Arg Arg Asp Val Gln Val Ser Asp Arg Val 710 715 720	1056
CGC AGC ATT GCC AAG GAC GTG GCC CTG GCC TGC AAG ACC TCT GAG ACG Arg Ser Ile Ala Lys Asp Val Ala Leu Ala Cys Lys Thr Ser Glu Thr 725 730 735	1104
GTA CCG CGG CCC GCA TGACTAGGCG TGGACCTGCC CATGGTGCCT GTCAGCCCGC Val Pro Arg Pro Ala 740	1159
AGAGCCCATC TACGCCAAC ACAGAGCTCA CACAGGTAC TGCTCTCTAG GCGGACACAC CCTGGGCCCT GAGCATCCAG AGCCTGGAT GGGCTTTCC CTGTGGCCA GGGATGCTCG GTCCCAGAGG AGGACCTAGT GACATCATGG GACAGGTCAA AGCATTAGGG CCACCTCCAT GGCCCCAGAC AGACTAAAGC TGCCCTCCTG GTGCAGGGCC GAGGGGACAC AAGGACCTAC CTGGAAGCAG CTGACATGCT GGTGGACGGC CGTTACTGGA GCCCGTGCCTC CTCCCTCCCC GTGCTTCATG TGACTCTTGG CCTCTCTGCT GCTGCGTTGG CAGAACCTG GGTGGGCAGG CACCCGGAGG AGGAGCAGCA GCTGTGTCAT CCTGTGCCCT CCATGTGCTG TGTGCTGTTT GCATGGCAGG GCTCCAGCTG CCTTCAGCCC TGTGACGTCT CCTCAGGGCA GCTGGACAGG CTTGGCACGG CCCGGGAAGT GCAGCAGGCA GCTTTCTTT GGGGTGGGAC TTGCCCTGAG CTTGGAGCTG CCACCTGGAG GACTTGCCTG TTCCGACTCC ACCTGTGCAG CCGGGGCCAC CCCAGGAGAA AGTGTCCAGG TGGGGCTGG CAGTCCCTGG CTGCAG	1219 1279 1339 1399 1459 1519 1579 1639 1699 1759 1805

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 370 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Glu Pro Leu Phe Pro Ala Pro Phe Trp Glu Val Ile Tyr Gly Ser
1 5 10 15

His Leu Gln Gly Asn Leu Ser Leu Leu Ser Pro Asn His Ser Leu Leu
20 25 30

Pro Pro His Leu Leu Asn Ala Ser His Gly Ala Phe Leu Pro Leu
35 40 45

Gly Leu Lys Val Thr Ile Val Gly Leu Tyr Leu Ala Val Cys Val Gly
50 55 60

Gly Leu Leu Gly Asn Cys Leu Val Met Tyr Val Ile Leu Arg His Thr
65 70 75 80

Lys Met Lys Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu Ala Leu Ala
85 90 95

Asp Thr Leu Val Leu Leu Thr Leu Pro Phe Gln Gly Thr Asp Ile Leu
100 105 110

Leu Gly Phe Trp Pro Phe Gly Asn Ala Leu Cys Lys Thr Val Ile Ala
115 120 125

Ile Asp Tyr Tyr Asn Met Phe Thr Ser Thr Phe Thr Leu Thr Ala Met
130 135 140

Ser Val Asp Arg Tyr Val Ala Ile Cys His Pro Ile Arg Ala Leu Asp
145 150 155 160

Val Arg Thr Ser Ser Lys Ala Gln Ala Val Asn Val Ala Ile Trp Ala
165 170 175

Leu Ala Ser Val Val Gly Val Pro Val Ala Ile Met Gly Ser Ala Gln
180 185 190

Val Glu Asp Glu Glu Ile Glu Cys Leu Val Glu Ile Pro Thr Pro Gln
195 200 205

Asp Tyr Trp Gly Pro Val Phe Ala Ile Cys Ile Phe Leu Phe Ser Phe
210 215 220

Ile Val Pro Val Leu Val Ile Ser Val Cys Tyr Ser Leu Met Ile Arg
225 230 235 240

Arg Leu Arg Gly Val Arg Leu Leu Ser Gly Ser Arg Glu Lys Asp Arg
245 250 255

Asn Leu Arg Arg Ile Thr Arg Leu Val Leu Val Val Ala Val Phe
260 265 270

Val Gly Cys Trp Thr Pro Val Gln Val Phe Val Leu Ala Gln Gly Leu
275 280 285

Gly Val Gln Pro Ser Ser Glu Thr Ala Val Ala Ile Leu Arg Phe Cys
290 295 300

Thr Ala Leu Gly Tyr Val Asn Ser Cys Leu Asn Pro Ile Leu Tyr Ala
305 310 315 320

Phe Leu Asp Glu Asn Phe Lys Ala Cys Phe Arg Lys Phe Cys Cys Ala
325 330 335

Ser Ala Leu Arg Arg Asp Val Gln Val Ser Asp Arg Val Arg Ser Ile
340 345 350

Ala Lys Asp Val Ala Leu Ala Cys Lys Thr Ser Glu Thr Val Pro Arg
355 360 365

Pro Ala
370

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 415 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: group(9, 12, 33, 40, 48)
- (D) OTHER INFORMATION: /note= "Extracellular Asn residues
that are consensus sites for N-linked glycosylation"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Asp Ser Ser Ala Ala Pro Thr Asn Ala Ser Asn Cys Thr Asp Ala
1 5 10 15

Leu Ala Tyr Ser Ser Cys Ser Pro Ala Pro Ser Pro Gly Ser Trp Val
20 25 30

Asn Leu Ser His Leu Asp Gly Asn Leu Ser Asp Pro Cys Gly Pro Asn
35 40 45

Arg Thr Asn Leu Gly Gly Arg Asp Ser Leu Cys Pro Pro Thr Gly Ser
50 55 60

Pro Ser Met Ile Thr Ala Ile Thr Ile Met Ala Leu Tyr Ser Ile Val
65 70 75 80

Cys Val Val Gly Leu Phe Gly Asn Phe Leu Val Met Tyr Val Ile Val
85 90 95

Arg Tyr Thr Lys Met Lys Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu
100 105 110

Ala Leu Ala Asp Ala Leu Ala Thr Ser Thr Leu Pro Phe Gln Ser Val
115 120 125

Asn Tyr Leu Met Gly Thr Trp Pro Phe Gly Thr Ile Leu Cys Lys Ile
130 135 140

Val Ile Ser Ile Asp Tyr Tyr Asn Met Phe Thr Ser Ile Phe Thr Leu
145 150 155 160

Cys Thr Met Ser Val Asp Arg Tyr Ile Ala Val Cys His Pro Val Lys
165 170 175

Ala Leu Asp Phe Arg Thr Pro Arg Asn Ala Lys Ile Ile Asn Val Cys
180 185 190

Asn Trp Ile Leu Ser Ser Ala Ile Gly Leu Pro Val Met Phe Met Ala
195 200 205

Thr Thr Lys Tyr Arg Gln Gly Ser Ile Asp Cys Thr Leu Thr Phe Ser
210 215 220

His Pro Thr Trp Tyr Trp Glu Asn Leu Val Lys Ile Cys Val Phe Ile
225 230 235 240

Phe Ala Phe Ile Met Pro Val Leu Ile Ile Thr Val Cys Tyr Gly Leu
245 250 255

Met Ile Leu Arg Leu Lys Ser Val Arg Met Leu Ser Gly Ser Lys Glu
260 265 270

Lys Asp Arg Asn Leu Arg Arg Ile Thr Arg Met Val Leu Val Val Val
275 280 285

Ala Val Phe Ile Val Cys Trp Thr Pro Ile His Ile Tyr Val Ile Ile
290 295 300

Lys Ala Leu Val Thr Ile Pro Glu Thr Thr Phe Gln Thr Val Ser Trp
305 310 315 320

His Phe Cys Ile Ala Leu Gly Tyr Thr Asn Ser Cys Leu Asn Pro Val
325 330 335

Leu Tyr Ala Phe Leu Asp Glu Asn Phe Lys Arg Cys Phe Arg Glu Phe
340 345 350

Cys Ile Pro Thr Ser Ser Asn Ile Glu Gln Gln Asn Ser Thr Arg Ile
355 360 365

Arg Gln Asn Thr Arg Asp His Pro Ser Thr Ala Asn Thr Val Asp Arg
370 375 380

Thr Asn His Gln Leu Glu Asn Leu Glu Ala Glu Thr Ala Pro Leu Pro
385 390 395 400

Phe Ile Gly Phe Ile Arg Ser Thr Ser Glu Gln Glu Asn Cys Glu
405 410 415

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 370 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: group(18, 33)
- (D) OTHER INFORMATION: /note= "Extracellular Asn residues that are consensus sites for N-linked glycosylation"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Glu Pro Ala Pro Ser Ala Gly Ala Glu Leu Gln Pro Pro Leu Phe
1 5 10 15

Ala Asn Ala Ser Asp Ala Tyr Pro Ser Ala Cys Pro Ser Ala Gly Ala
20 25 30

Asn Ala Ser Gly Pro Pro Ala Arg Ser Ala Ser Ser Leu Ala Leu Ala
35 40 45

Ile Ala Ile Thr Ala Leu Tyr Ser Ala Val Cys Ala Val Gly Leu Ile
50 55 60

Gly Asn Val Leu Val Met Phe Gly Ile Val Arg Tyr Thr Lys Met Lys
65 70 75 80

Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu Ala Leu Ala Asp Ala Leu
85 90 95

Ala Thr Ser Thr Leu Pro Phe Gln Ser Ala Lys Tyr Leu Met Glu Thr
100 105 110

Trp Pro Phe Gly Glu Leu Leu Cys Lys Ala Val Leu Ser Ile Asp Tyr
115 120 125

Tyr Asn Met Phe Thr Ser Ile Phe Thr Leu Thr Met Met Ser Val Asp
130 135 140

Arg Tyr Ile Ala Val Cys His Pro Val Lys Ala Leu Asp Phe Arg Thr
145 150 155 160

Pro Ala Lys Ala Lys Leu Ile Asn Ile Cys Ile Trp Val Leu Ala Ser
165 170 175

Gly Val Gly Val Pro Ile Met Val Met Ala Val Thr Arg Pro Arg Asp
180 185 190

Gly Ala Val Val Cys Met Leu Gln Phe Pro Ser Pro Ser Trp Tyr Trp
195 200 205

Asp Thr Val Thr Lys Ile Cys Val Phe Leu Phe Ala Phe Val Val Pro
210 215 220

Ile Leu Ile Ile Thr Val Cys Tyr Gly Leu Met Leu Leu Arg Leu Arg
225 230 235 240

Ser Val Arg Leu Leu Ser Gly Ser Lys Glu Lys Asp Arg Ser Leu Arg
245 250 255

Arg Ile Thr Arg Met Val Leu Val Val Gly Ala Phe Val Val Cys
260 265 270

Trp Ala Pro Ile His Ile Phe Val Ile Val Trp Thr Leu Val Asp Ile
275 280 285

Arg Arg Asp Pro Leu Val Val Ala Ala Leu His Leu Cys Ile Ala Leu
290 295 300

Gly Tyr Ala Asn Ser Ser Leu Asn Pro Val Leu Tyr Ala Phe Leu Asp
305 310 315 320

Glu Asn Phe Lys Arg Cys Phe Arg Gln Leu Cys Arg Lys Pro Cys Gly
325 330 335

Arg Pro Asp Pro Ser Ser Phe Ser Arg Ala Arg Glu Ala Thr Ala Arg
340 345 350

Glu Arg Val Thr Ala Cys Thr Pro Ser Asp Gly Pro Gly Gly Ala
355 360 365

Ala Ala
370

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 394 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: group(25, 39)
- (D) OTHER INFORMATION: /note= "Extracellular Asn residues that are consensus sites for N-linked glycosylation"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Glu Ser Pro Ile Gln Ile Phe Arg Gly Asp Pro Gly Pro Thr Cys
1 5 10 15

Ser Pro Ser Ala Cys Leu Leu Pro Asn Ser Ser Ser Trp Phe Pro Asn
20 25 30

Trp Ala Glu Ser Asp Ser Asn Gly Ser Val Gly Ser Glu Asp Gln Gln
35 40 45

Leu Glu Ser Ala His Ile Ser Pro Ala Ile Pro Val Ile Ile Thr Ala
50 55 60

Val Tyr Ser Val Val Phe Val Val Gly Leu Val Gly Asn Ser Leu Val
65 70 75 80

Met Phe Val Ile Ile Arg Tyr Thr Lys Met Lys Thr Ala Thr Asn Ile
85 90 95

Tyr Ile Phe Asn Leu Ala Leu Ala Asp Ala Leu Val Thr Thr Thr Met
100 105 110

Pro Phe Gln Ser Ala Val Tyr Leu Met Asn Ser Trp Pro Phe Gly Asp
115 120 125

Val Leu Cys Lys Ile Val Ile Ser Ile Asp Tyr Tyr Asn Met Phe Thr
130 135 140

Ser Ile Phe Thr Leu Thr Met Met Ser Val Asp Arg Tyr Ile Ala Val
145 150 155 160

Cys His Pro Val Lys Ala Leu Asp Phe Arg Thr Pro Leu Lys Ala Lys
165 170 175

Ile Ile Asn Ile Cys Ile Trp Leu Leu Ala Ser Ser Val Gly Ile Ser
180 185 190

Ala Ile Val Leu Gly Gly Thr Lys Val Arg Glu Asp Val Ile Glu Cys
195 200 205

Ser Leu Gln Phe Pro Asp Asp Glu Trp Trp Asp Leu Phe Met Lys Ile
210 215 220

Cys Val Phe Val Phe Ala Phe Val Ile Pro Val Leu Ile Ile Ile Val
225 230 235 240

Cys Tyr Thr Leu Met Ile Leu Arg Leu Lys Ser Val Arg Leu Leu Ser
245 250 255

Gly Ser Arg Glu Lys Asp Arg Asn Leu Arg Arg Ile Thr Lys Leu Val
260 265 270

Leu Val Val Val Ala Val Phe Ile Ile Cys Trp Thr Pro Ile His Ile
275 280 285

Phe Ile Leu Val Glu Ala Leu Gly Ser Thr Ser His Ser Thr Ala Ala
290 295 300

Leu Ser Ser Tyr Tyr Phe Cys Ile Ala Leu Gly Tyr Thr Asn Ser Ser
305 310 315 320

Leu Asn Pro Val Leu Tyr Ala Phe Leu Asp Glu Asn Phe Lys Arg Cys
325 330 335

Phe Arg Asp Phe Cys Phe Pro Ile Lys Met Arg Met Glu Arg Gln Ser
340 345 350

Thr Asn Arg Val Arg Asn Thr Val Gln Asp Pro Ala Ser Met Arg Asp
355 360 365

Val Gly Gly Met Asn Lys Pro Val Glu Gln Met His Glu Lys Ser Ile
370 375 380

Pro Tyr Ser Gln Glu Thr Leu Val Val Asp
385 390

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 367 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: group(19, 26, 37)
- (C) OTHER INFORMATION: /note= "Extracellular Asn residues that are consensus sites for N-linked glycosylation"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Glu Pro Leu Phe Pro Ala Pro Phe Trp Ile Tyr Gly Ser His Leu
1 5 10 15

Gln Gly Asn Leu Ser Leu Leu Ser Pro Asn His Ser Leu Leu Pro Pro
20 25 30

His Leu Leu Leu Asn Ala Ser His Gly Ala Leu Pro Leu Gly Leu Lys
35 40 45

Val Thr Ile Val Gly Leu Tyr Leu Ala Val Cys Val Gly Gly Leu Leu
50 55 60

Gly Asn Cys Leu Val Met Tyr Val Leu Leu Arg His Thr Lys Met Lys
65 70 75 80

Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu Ala Leu Ala Asp Thr Leu
85 90 95

Val Leu Leu Thr Leu Pro Phe Gln Gly Thr Asp Ile Leu Leu Gly Pro
100 105 110

Trp Pro Phe Gly Asn Ala Leu Cys Lys Thr Val Ile Ala Ile Asp Tyr
115 120 125

Tyr Asn Met Phe Thr Ser Thr Phe Thr Leu Thr Ala Met Ser Val Asp
130 135 140

Arg Tyr Val Ala Ile Cys His Pro Ile Arg Ala Leu Asp Val Arg Thr
145 150 155 160

Ser Ser Lys Ala Gln Ala Val Asn Val Ala Ile Trp Ala Leu Ala Ser
165 170 175

Val Val Gly Val Pro Val Ala Ile Met Gly Ser Ala Gln Val Glu Asp
180 185 190

Glu Glu Ile Glu Cys Leu Val Glu Ile Pro Thr Pro Gln Asp Tyr Trp
195 200 205

Gly Pro Val Phe Ala Ile Cys Ile Phe Leu Phe Ser Phe Ile Val Pro
210 215 220

Val Leu Val Ile Ser Val Cys Tyr Ser Leu Met Ile Arg Arg Leu Arg
225 230 235 240

Gly Val Arg Leu Leu Ser Gly Ser Arg Glu Lys Asp Arg Asn Leu Arg
245 250 255

Arg Ile Thr Arg Leu Val Leu Val Val Val Ala Val Phe Val Gly Cys
260 265 270

Trp Thr Pro Val Gln Val Phe Val Leu Ala Gln Gly Leu Gly Val Gln
275 280 285

Pro Ser Ser Glu Thr Ala Val Ala Ile Leu Arg Phe Cys Thr Ala Leu
290 295 300

Gly Tyr Val Asn Ser Cys Leu Asn Pro Ile Leu Tyr Ala Phe Leu Asp
305 310 315 320

Glu Asn Phe Lys Ala Cys Phe Arg Lys Phe Cys Cys Ala Ser Ala Leu
325 330 335

Arg Arg Asp Val Gln Val Ser Asp Arg Val Arg Ser Ile Ala Lys Asp
340 345 350

Val Ala Leu Ala Cys Lys Thr Ser Glu Thr Val Pro Arg Pro Ala
355 360 365

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 15
- (D) OTHER INFORMATION: /note= "Extracellular Asn residue which is a consensus site for N-linked glycosylation"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Glu Glu Gly Gly Asp Phe Asp Asn Tyr Tyr Gly Ala Asp Asn Gln
1 5 10 15

Ser Glu Cys Glu Tyr Thr Asp Trp Lys Ser Ser Gly Ala Leu Ile Pro
20 25 30

Ala Ile Tyr Met Leu Val Phe Leu Leu Gly Thr Thr Gly Asn Gly Leu
35 40 45

Val Leu Trp Thr Val Phe Arg Ser Ser Arg Lys Arg Arg Ser Ala Asp
50 55 60

Ile Phe Ile Ala Ser Leu Ala Val Ala Asp Leu Thr Phe Val Val Thr
65 70 75 80

Leu Pro Leu Ala Thr Tyr Thr Arg Asp Tyr Asp Trp Pro Phe Gly
85 90 95

Thr Phe Phe Cys Lys Leu Ser Ser Tyr Leu Ile Phe Val Asn Met Tyr
100 105 110

Ala Ser Val Phe Cys Leu Thr Gly Leu Ser Phe Asp Arg Tyr Leu Ala
115 120 125

Ile Val Arg Pro Val Ala Asn Ala Arg Leu Arg Leu Arg Val Ser Gly
130 135 140

Ala Val Ala Thr Ala Val Leu Trp Val Leu Ala Ala Leu Leu Ala Met
145 150 155 160

Pro Val Met Val Leu Thr Thr Gly Asp Leu Glu Asn Thr Thr Val Gln
165 170 175

Cys Tyr Met Asp Tyr Ser Ser Ser Glu Trp Ala Trp Glu Val Gly Leu
180 185 190

Gly Val Ser Ser Thr Thr Val Gly Phe Val Val Pro Phe Thr Ile Met
195 200 205

Leu Thr Cys Tyr Phe Phe Ile Ala Gln Thr Ile Ala Gly His Phe Arg
210 215 220

Lys Glu Arg Ile Glu Gly Leu Arg Lys Arg Arg Arg Leu Leu Ser Ile
225 230 235 240

Ile Val Val Leu Val Val Thr Phe Ala Leu Cys Trp Met Pro Tyr His
245 250 255

Leu Tyr Met Leu Gly Ser Leu Leu His Trp Pro Cys Asp Asp Leu Phe
260 265 270

Leu Met Asn Ile Phe Pro Tyr Cys Thr Cys Ile Ser Tyr Val Asn Ser
275 280 285

Cys Leu Asn Pro Phe Leu Tyr Ala Phe Phe Asp Pro Arg Phe Arg Ala
290 295 300

Cys Thr Ser Met Leu Cys Cys Gly Gln Ser Arg Cys Ala Gly Thr Ser
305 310 315 320

His Ser Ser Ser Gly Glu Lys Ser Ala Ser Tyr Ser Ser Gly His Ser
325 330 335

Gln Gly Pro Gly Pro Asn Met Gly Lys Gly Gly
340 345

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5
- (D) OTHER INFORMATION: /note= "Any amino acid"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Tyr Gly Gly Phe Xaa
1 5